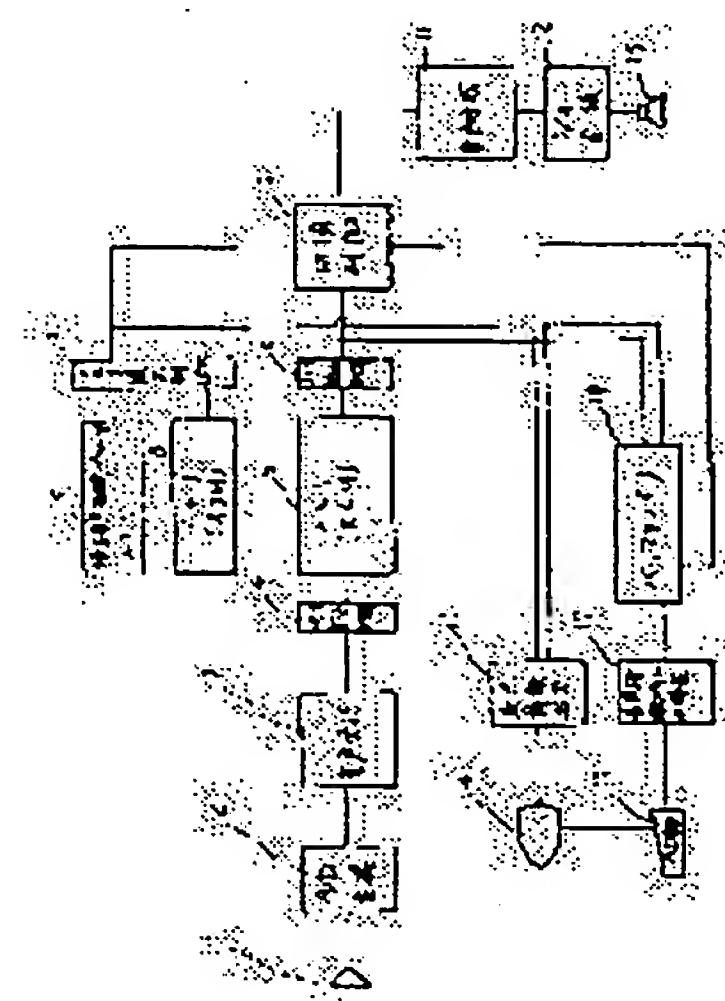


## PATENT ABSTRACTS OF JAPAN

(11) Publication number : **61-006732**(43) Date of publication of application : **13.01.1986**

(51) Int.Cl.

**G06F 3/16****G09B 5/04****G09B 17/00****G10L 3/00**(21) Application number : **59-127809**(71) Applicant : **FUJITSU LTD**(22) Date of filing : **21.06.1984**(72) Inventor : **KOBAYASHI ATSUSHI****(54) VOCAL TRAINING DEVICE****(57) Abstract:**

**PURPOSE:** To execute vocal training actively by storing voice feature parameters of a model and a vocal training person and displaying these parameters in parallel every vocal training.

**CONSTITUTION:** A vocal training device is constituted of a microphone 1 for inputting a vocal sound, an A/D converter 2 for converting the inputted vocal sound into digital data, a voice analyzing circuit 3 for converting the digital data into a voice feature parameter time series, a memory writing control circuit 4 for the time series, a display part 14, a voice synthesizing circuit 11, a speaker 13, etc. A vocal sound of the vocal training person which is inputted through said microphone 1 is converted into a voice feature parameter time series and the converted data are stored in a RAM5 and displayed in parallel with the voice feature of the model on the display part 14. The vocal training person repeats vocal training on the basis of the displayed information and corrects the accent, intonation and rhythm of his voice, so that the degree of training effects can be grasped by himself.

## Sole Claim:

A vocal training device for practicing vocal training by oneself based on model voice data displayed as visible data or reproduced as a voice, comprising model voice data storing means which have stored a predetermined vocal text as model voice feature parameters; voice analyzing means for analyzing a feature of data of a voice uttered by a vocal training person; accumulating means for accumulating and storing voice feature parameters of the vocal training person, which are outputs from the voice analyzing means; feature parameter modifying means for modifying accumulated and stored voice feature parameters of the model voice data and the vocal training person's voice data; and voice synthesizing means for synthesizing a voice from accumulated and stored voice feature parameters of the model voice data and the vocal training person's voice data.

## Detailed Description of the Invention:

## (1) (Stage before listening to model voice data)

For predetermined text items for vocal training, a vocal sound to be a model is analyzed, converted to a voice feature parameter time series and stored in the memory (ROM) 8.

As the voice analysis/synthesis method used here, there are various methods. In this embodiment, it is assumed that the PARCOR analysis/synthesis method commonly used is to be used.

The memory reading control circuit 9 reads voice feature parameters from this memory (ROM) 8 and outputs them to the voice synthesizing circuit 11 via the switching control circuit 10. The voice synthesizing circuit 11 synthesizes and creates voice for items to train based on the voice feature parameters, and reproduces it from the speaker 13 to have a vocal training person listen to it.

It is possible to modify a voice feature parameter time series and synthesize multiple voices as shown in Figure 2 to have the vocal training person listen to it. The voice feature parameters to be modified are the vocal code information parameter and the amplitude parameter.

The vocal code information parameter corresponds to the accent and intonation of an uttered voice, and the amplitude parameter corresponds to the strength and weakness of the uttered voice.

## (2) (Stage at which training is performed based on the model voice data, and the effect of the training is checked)

The vocal training person performs vocal training based on the synthesized vocal sound to be a model in accordance with instructions, inputs his voice from the microphone 1 into the voice analyzing circuit 3 via the A/D converter 2, performs voice analysis with the PARCOR analysis/synthesis method described in (1) to make conversion to a voice feature parameter time series, and stores the voice feature parameter time series in the memory (RAM) 5 via the memory writing control circuit 4.

In this case, the voice features of the model voice and the vocal training person's voice are displayed in parallel on the display part 14, as shown in Figure 3. The voice cord information parameter time series and the amplitude parameter time series are used as the data to be displayed.

In Figure 3, reference numerals (1), (1)', (2) and (2)' denote the amplitude information about the model voice, the voice cord information about the model voice, the amplitude information about the vocal training person, and the voice cord information about the vocal training person, respectively.

The vocal training person repeats vocal training based on the displayed information. For example, by comparing and checking the displayed information, he tries to utter voice by paying attention to the accent or intonation to bring the voice cord information (2)' close to the model voice and tries to utter voice by paying attention to the strength and weakness of voice to bring the

amplitude information (1)' close to the model voice. As for to what degree the accent, intonation, and strength and weakness should be specifically modified, description will be made below about the modification to be made for each of the displayed information.

That is, for example, in order to modify the voice cord information (2)' about the vocal training person, time T to be modified (all the horizontal axes in Figures 3 and 4 indicate time) and a modification degree n (this is a quantized modification degree, and it corresponds to the displayed scale value and also corresponds to a modification degree inputted from an input part 15) are specified and inputted from the input part 15, and thereby the parameter of the time T of the voice cord information parameter time series is modified by the modification degree n as shown in Figure 4.

In this case, the voice cord information parameters t hours before and after the time T are also corrected. There are also various methods for making correction for t hours before and after the time T. In this embodiment, however, it is assumed that linear correction is made, that is, the parameters are linearly corrected.

By performing the modification operation described above, it is possible to know to what degree the accent, intonation, and strength and weakness should be modified.

According to the present invention described above, there is an advantage that a vocal training person can perform vocal training appropriate for his ability by grasping the direction of the vocal training and the degree of the effect of the training himself.

⑨ 日本国特許庁 (JP) ⑩ 特許出願公開  
⑪ 公開特許公報 (A) 昭61-6732

⑫ Int. Cl. 4  
G 06 F 3/16  
G 09 B 5/04  
17/00  
G 10 L 3/00

識別記号 廈内整理番号  
7622-5B  
6612-2C  
6612-2C  
7350-5D

⑬ 公開 昭和61年(1986)1月13日  
審査請求 未請求 発明の数 1 (全5頁)

⑭ 発明の名称 発声訓練装置

⑮ 特願 昭59-127809  
⑯ 出願 昭59(1984)6月21日

⑰ 発明者 小林 敦仁 川崎市中原区上小田中1015番地 富士通株式会社内  
⑱ 出願人 富士通株式会社 川崎市中原区上小田中1015番地  
⑲ 代理人 弁理士 松岡 宏四郎

明細書

1. 発明の名称

発声訓練装置

2. 特許請求の範囲

可視データとして表示され、あるいは音声として再生された手本音声データをもとに発声訓練を独習する装置において、予め定められた発声テキストを手本音声特徴パラメータとして収録している手本音声データ記憶手段と、発声訓練者が発声した音声データの特徴を分析する音声分析手段と、前記音声分析手段の出力である発声訓練者音声特徴パラメータを蓄積格納する蓄積手段と、蓄積格納された手本音声データ及び発声訓練者音声データの音声特徴パラメータを修正する特徴パラメータ修正手段と、蓄積格納された手本音声データ及び発声訓練者音声データの音声特徴パラメータから音声を合成する音声合成手段とを設けたことを特徴とする発声訓練装置。

3. 発明の詳細な説明

(産業上の利用分野)

本発明は、手本音声データ及び発声訓練者音声データを音声特徴パラメータに分析格納し、前記音声特徴パラメータを修正表示し、合成音声を聞くことにより発声訓練を行う発声訓練装置に関する。

発声訓練装置は学校で使用されるような大勢を対象とする発声訓練装置や、個人の独習を目的とする発声訓練装置等その利用形態に応じた各種タイプが実用化されている。

しかし、語学訓練、アクセントの矯正、聴話者の発声訓練と目的に応じた各種タイプの発声訓練装置に対する要望は多く、特に個人の独習を目的とする発声訓練装置においては効率的な機能構成で、しかも能動的な学習が可能な発声訓練装置の開発実用化が要望されている。

〔従来の技術〕

従来の発声訓練装置、特に語学等の独習を目的とする発声訓練装置として圧倒的に多いのは、磁気テープ等に記録された手本音声データを再生し、その音声データを聞いて発声訓練者が発声するこ

とにより訓練する方法である。

この方法は、発声訓練者の発声の仕方が正しくなされているかどうかの判断は、発声訓練者自身の主観的な判断による場合が多い。

又最近は、パーソナルコンピュータを利用した発声訓練装置で音声データの特徴を抽出して表示することにより学習する装置も開発されている。

〔発明が解決しようとする問題点〕

しかし、上記のような発声訓練者自身が主観的に成果を評価判断する従来の発声訓練装置では、もし発声訓練者自身の判断が間違っているか、又は不充分の場合は正しい発声の仕方がなされない状態で訓練されてしまうと言う問題点がある。

又音声データの特徴を抽出して表示する装置でも発声訓練者は常に受け身の立場にあり、長期間の発声訓練には耐え難いと言う問題点がある。

〔問題点を解決するための手段〕

本発明は、上記問題点を解消した新規な発声訓練装置を実現することを目的とするものであり、該問題点は、予め定められた発声テキストを手本

音声特徴パラメータとして収録している手本音声データ記憶手段と、発声訓練者が発声した音声データの特徴を分析する音声分析手段と、前記音声分析手段の出力である発声訓練者音声特徴パラメータを蓄積格納する蓄積手段と、蓄積格納された手本音声データ及び発声訓練者音声データの音声特徴パラメータを修正する特徴パラメータ修正手段と、蓄積格納された手本音声データ及び発声訓練者音声データの音声特徴パラメータから音声を合成する音声合成手段とを設けてなる本発明による発声訓練装置により解決される。

〔作用〕

即ち、予め定められた発声訓練用テキスト項目について手本となる音声データを音声分析し、その音声特徴パラメータを抽出し、音声特徴パラメータを蓄積格納して置く。

同様に発声訓練者の発声した音声データも音声特徴パラメータを抽出し、その音声特徴パラメータを蓄積格納して置き、発声訓練の都度並列表示される音声特徴パラメータを見ながら手本又は発

声訓練者の音声特徴パラメータを修正することにより、発声訓練のポイントを探求しながら能動的に発声訓練が出来るようにしたものである。

〔実施例〕

以下本発明の要旨を第1図に示す実施例により具体的に説明する。

第1図は本発明に係る発声訓練装置を示すブロックダイヤグラム、第2図は手本の発声の声帯情報パラメータの修正状況、第3図は手本の発声と発声訓練者の発声との声帯情報パラメータの比較表示状況、第4図は発声訓練者の発声の声帯情報パラメータの修正状況をそれぞれ示す。

本実施例は、

予め定められた発声訓練用テキストに沿って発声された発声訓練者の発聲音を入力するマイクロホン1。

マイクロホン1から入力されたアナログ発聲音をディジタルに変換するアナログ/ディジタル変換器(以下A/D変換器と称する)2。

ディジタルデータに変換された声帯情報を音声分

析(バーコール音声分析)し音声特徴パラメータ時系列に変換する音声分析回路3。

音声特徴パラメータ時系列を所定メモリへ蓄込むための制御をするメモリ書き込み制御回路4。

発声訓練者の音声特徴パラメータ時系列を記憶するメモリ(RAM)5。

メモリ(RAM)5の内容の読出し制御を行うメモリ読出制御回路6。

手本用の修正音声特徴パラメータを記憶しているメモリ7。

手本の音声特徴パラメータ時系列を記憶しているメモリ(ROM)8。

メモリ7、メモリ(ROM)8の内容の読出し制御を行うメモリ読出制御回路9。

各種メモリ(5,7,8,18)からの出力の切り換えを行う切換制御回路10。

切換制御回路10から出力される音声特徴パラメータから各種の音声を合成する音声合成回路11。

音声合成回路11のディジタル出力をアナログに変換するディジタル/アナログ変換器(D/A変換

### 特開昭61-6732(3)

器) 12,  
D/A変換器12の出力を音声として再生するスピーカ13,  
表示情報出力回路16から出力される声帯情報パラメータ時系列と振幅情報パラメータ時系列とを並列表示する表示部14,  
音声特徴パラメータを修正するためのデータを入力指示する入力部15,  
メモリ読出制御回路6, 9により読み出された手本用及び発声訓練者用音声特徴パラメータのうち声帯情報パラメータ時系列と振幅情報パラメータ時系列とに変換出力する表示情報出力回路16,  
入力部15の指示するデータにより音声特徴パラメータを修正する音声特徴パラメータ修正部17,  
声帯情報パラメータ及び振幅情報パラメータからなる音声特徴パラメータを一時的に格納するバッファメモリ18,  
から構成されている。

次に本実施例の動作を。

#### (1) 手本音声データを聞くまでの段階:

(2) 手本音声データをもとにして訓練し訓練結果を確認する段階。  
に分けて説明する。

##### (1) (手本音声データを聞くまでの段階)

予め定められた発声訓練用デキスト項目について、手本となる発声の音声を音声分析し音声特徴パラメータの時系列に変換してメモリ(ROM)8に格納して置く。尚メモリ(ROM)8へデータを格納する動作は図示していない装置で事前に行われているものとする。

更に、この時の音声分析合成方式としては、各種方式が存在するが、本実施例としては一般的に良く用いられるバーコール分析合成方式を用いた場合とする。又バーコール型の音声合成における音声特徴パラメータは、

- ・バーコール係数、
- ・声帯情報パラメータ、
- ・有声・無声判定パラメータ、
- ・振幅パラメータ、

である。従って、メモリ(ROM)8には上記4つ

のパラメータをそれぞれ専用に格納するメモリ領域が設けられている。

このメモリ(ROM)8から音声特徴パラメータをメモリ読出制御回路9が読み出し、切換制御回路10経由音声合成回路11に出力する。音声合成回路11は音声特徴パラメータをもとに訓練する項目の音声を合成生成しスピーカ13で再生し、発声訓練者に聞かせる。

又音声特徴パラメータの時系列を修正して、第2図に示すように複数の発声を合成して発声訓練者に聞かせることが出来る。修正の対象となる音声特徴パラメータは声帯情報パラメータと振幅パラメータである。

尚声帯情報パラメータは発声音声のアクセント、イントネーションに対応するものであり、振幅パラメータは発声音声の強弱に対応するものである。又音声特徴パラメータの修正はこれら声帯情報パラメータと振幅パラメータとを修正することにより行い、修正データはメモリ7に格納されている修正音声特徴パラメータ(但し、修正声帯情報パラ

メータ、修正振幅パラメータとして格納している)を入力部15からの指定により取り出して行う。

声帯情報パラメータは発声音声の基本周波数であり、この基本周波数の時系列パターン形状を例えば第2図(A)の点線のように変形修正し、疑問文の形に合成出力することが出来る。又発声音声の振幅値の時系列パターンを例えば第2図(B)の点線のように変形修正し、強弱を変えた発声音声を合成出力することが出来る。

##### (2) (手本音声データをもとにして訓練し訓練結果を確認する段階)

発声訓練者は指示に従って手本となる発声の合成音をもとに発声訓練し、この音声をマイクロホン1よりA/D変換器2経由音声分析回路3に入力し、(1)項で説明したバーコール分析合成方式により音声分析し、音声特徴パラメータの時系列に変換し、この音声特徴パラメータの時系列をメモリ書き込み制御回路4によりメモリ(RAM)5に格納する。

ここで、第3図に示すように手本発声と発声訓

練習者発声の発声特徴を表示部14に並列表示する。表示するデータとしては、声帯情報パラメータ、振幅パラメータの時系列を用いる。尚表示データには目盛が付加されており、この目盛はデータを修正する場合等の単位として使用する。

又第3図において、①は手本発声の振幅情報、①'は手本発声の声帯情報、②は発声訓練者発声の振幅情報、②'は発声訓練者発声の声帯情報をそれぞれ示す。

発声訓練者は表示された情報をもとに発声訓練を繰り返す。例えば、表示された情報を比較検討し、声帯情報②'を手本発声に近づかせるには、アクセントやイントネーションを特に注意し発声し、振幅情報①'を手本発声に近づかせるには、発声の強弱を特に注意し発声する。尚具体的にどの程度アクセント、イントネーション及び強弱を修正するかは各表示情報に以下に説明するような修正を加えながら行う。

即ち、例えば、発声訓練者の声帯情報②'を修正するには修正すべき時刻T'（第3図、第4図の

横軸は全て時刻を示す）と修正度n（これは量子化された修正度であり、表示目盛と対応しており、又入力部15から入力する修正度とも対応している）を入力部15から指定入力することにより、声帯情報パラメータ時系列の時刻Tのパラメータを第4図に示すように修正度nだけ修正する。

尚この場合、時刻Tの前後 $\pm$ 時間内の声帯情報パラメータも補正する。この時刻Tの前後 $\pm$ 時間の補正方法も種々の方法があるが、本実施例では線形補正、即ち直線的にパラメータを補正するものとする。

以上の修正動作を行うことにより、どの程度アクセント、イントネーション及び強弱を修正すれば良いかが分かる。

又以上の方法とは逆に、手本発声の例えは声帯情報②を修正し、発声訓練者の声帯情報②'と一致するまでの修正データにより、発声訓練の仕方が正しい方向か否かや、後どの程度の発声訓練がどの方向で行えればよいか等の情報をつかみながら発声訓練を行うことも出来る。

#### （発明の効果）

以上のような本発明によれば、発声訓練の訓練の方向や訓練効果の程度を発声訓練者自身が把握しながら自分の能力にあった発声訓練が出来ると言う効果がある。

#### 4. 図面の簡単な説明

第1図は本発明に係る発声訓練装置を示すブロックダイヤグラム。

第2図は手本の発声の声帯情報パラメータの修正状況。

第3図は手本の発声と発声訓練者の発声との声帯情報パラメータの比較表示状況。

第4図は発声訓練者の発声の声帯情報パラメータの修正状況。

をそれぞれ示す。

図において、

- 1はマイクロホン、 2はA/D変換器、
- 3は音声分析回路、 4はメモリ書き込み制御回路
- 5はメモリ(RAM)、
- 6. 9はメモリ読み出し制御回路、

7は修正音声特徴パラメータメモリ、  
8はメモリ(ROM)、10は切換制御回路、  
11は音声合成回路、12はD/A変換回路、  
13はスピーカ、14は表示部、  
15は入力部、16は表示情報出力回路、  
17は音声特徴パラメータ修正回路、  
18はバッファメモリ、  
をそれぞれ示す。

代理人 弁理士 松岡宏四郎

Fig. 1

第 1 図

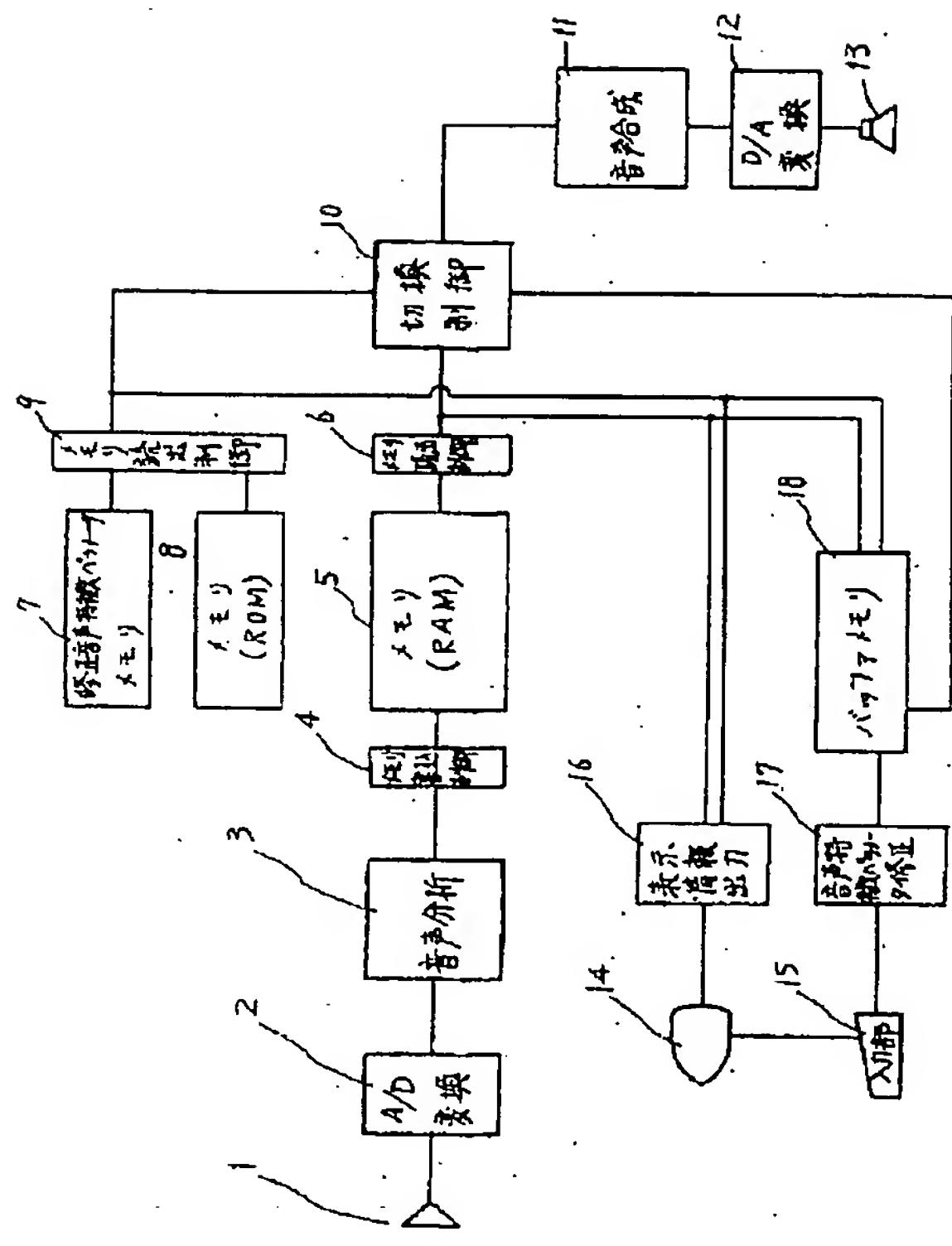


Fig. 2

特開昭61-6732(5)

第 2 図

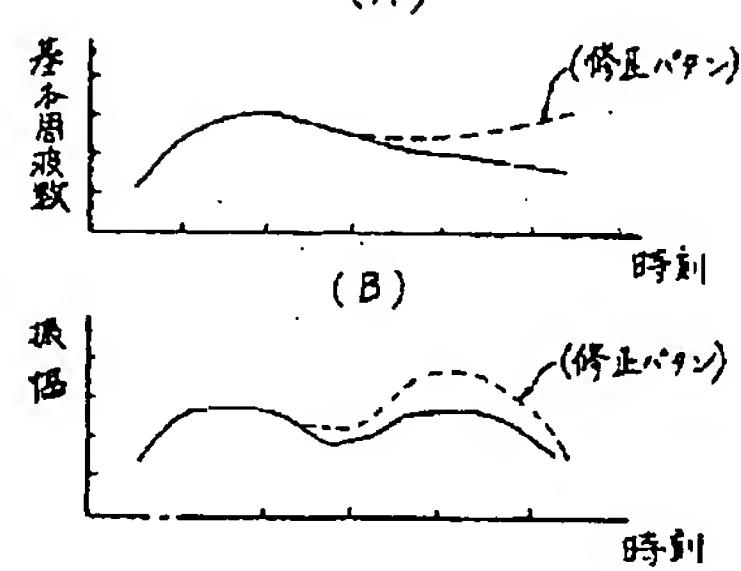


Fig. 3

第 3 図

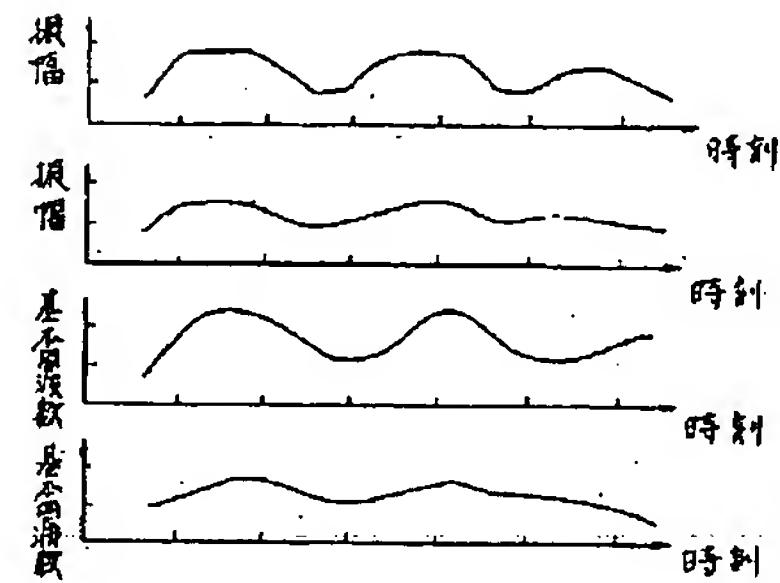
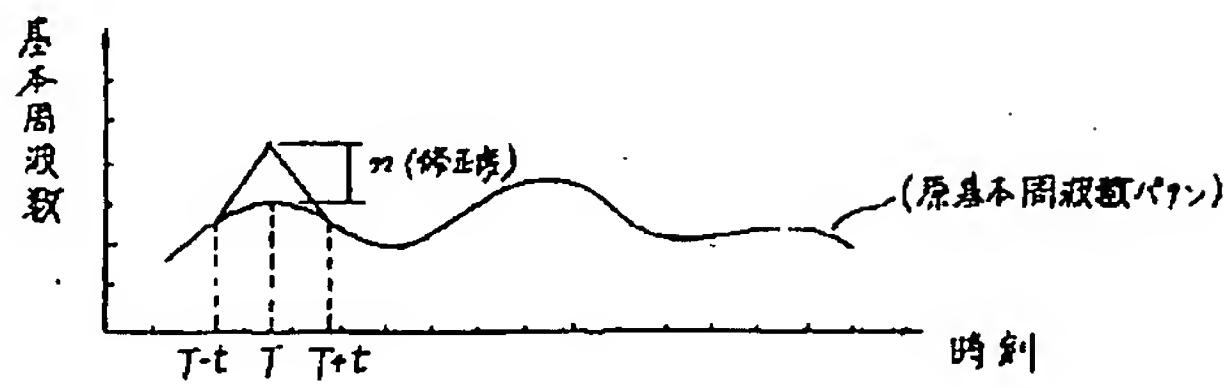


Fig. 4

第 4 図



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**